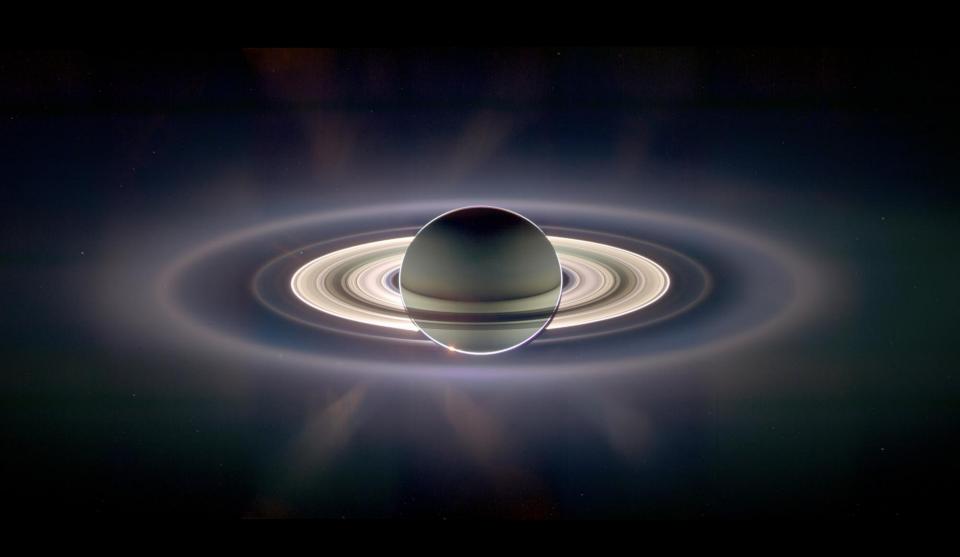
CC7220-1 LA WEB DE DATOS PRIMAVERA 2018

LECTURE 1: INTRODUCTION

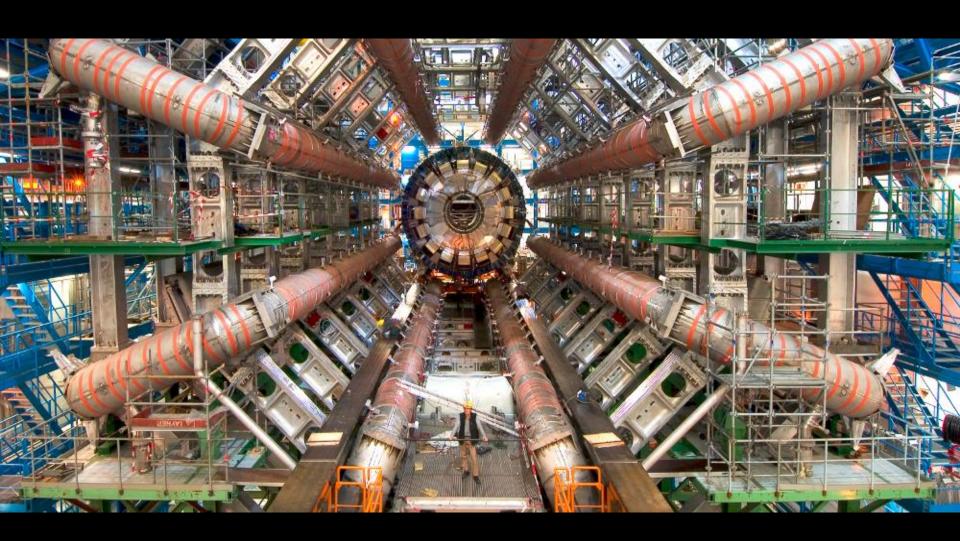
Aidan Hogan aidhog@gmail.com

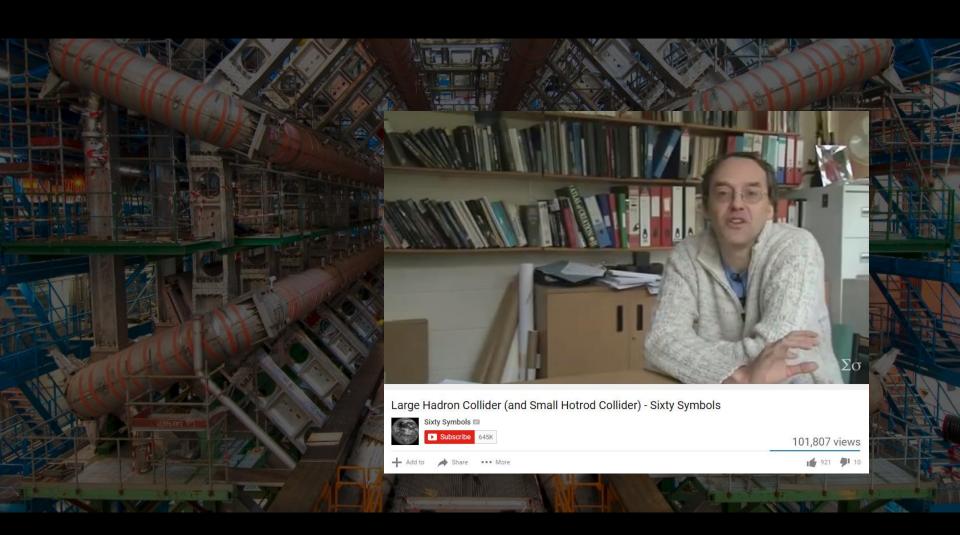


Who's this guy?









Videos about the symbols of physics and astronomy

"... is it worth spending this kind of money on these kinds of things?"

"... one answer is that fundamentally the whole human endeavour ..."

"... but the other thing you really have to give the particle physicists for [...] and that they should be more famous for [...] is that it was particle physicists who were entirely responsible for the World Wide Web ..."

"... they invented this thing [...] the World Wide Web [...] for disseminating their results amongst their colleagues [...] and so the hundreds of billions of pounds that the Internet is now worth comes from particle physicists messing around ..."

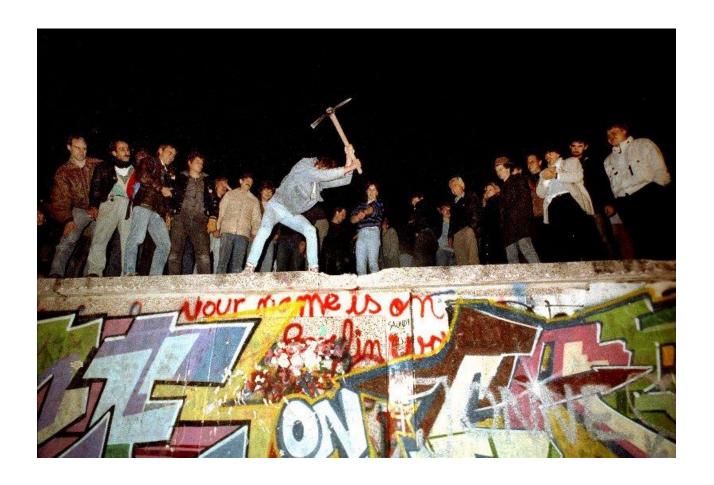
TIM BERNERS-LEE: PARTICLE PHYSICIST (APPARENTLY)



Who's this guy?

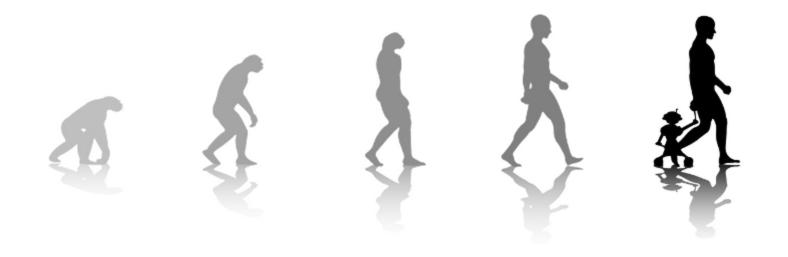
THE WEB

THE WEB IS NOW 29 YEARS OLD



How old is the Web?

THE FUTURE OF THE WEB?



What will the Web be like in 29 years time?

THE SEMANTIC WEB?

SEMANTIC WEB?



semantic web



Google Search

I'm Feeling Lucky

SEMANTIC WEB?

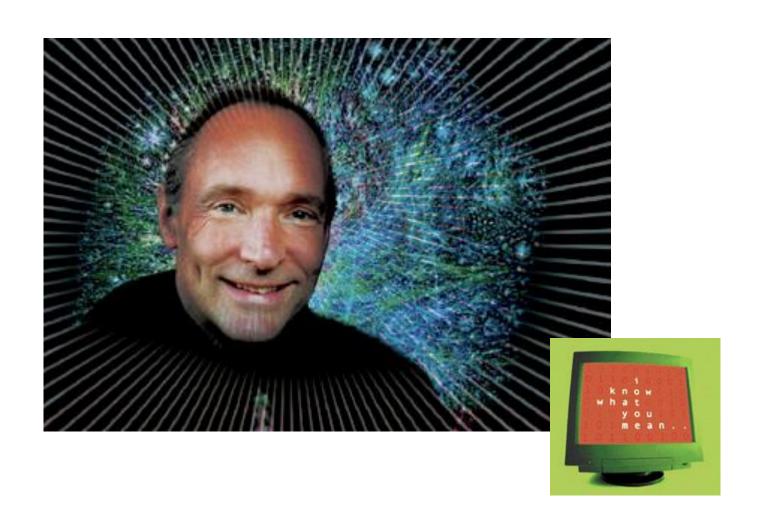


"The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated task as a sarry."



THE SEMANTIC WEB IS ...

... TIM BERNERS-LEE'S VISION OF THE FUTURE WEB?



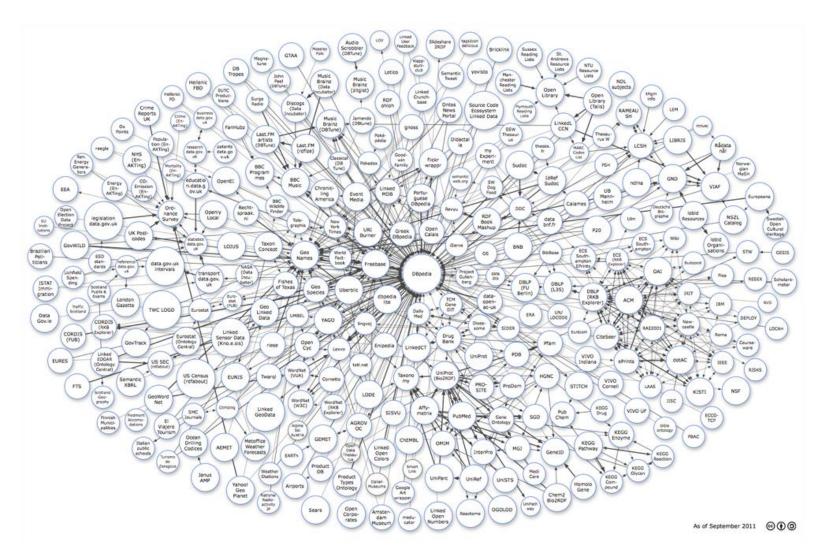
THE SEMANTIC WEB IS ...

... INTELLIGENT MACHINES MAKING SENSE OF THE WEB?

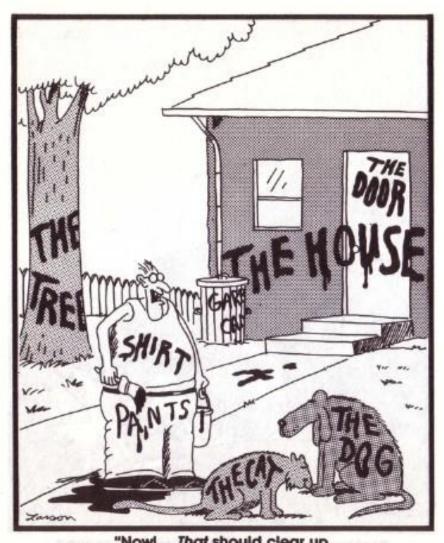


THE SEMANTIC WEB IS ...

... LINKING TOGETHER OPEN DATA?

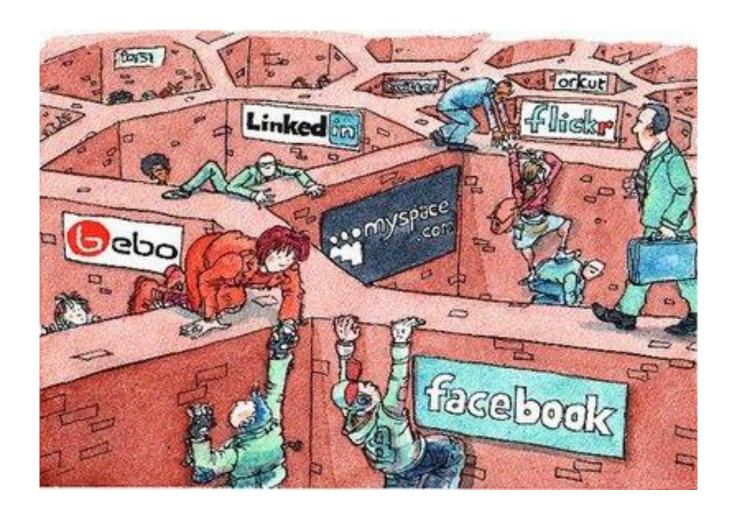


THE SEMANTIC WEB IS NAMING EVERYTHING?

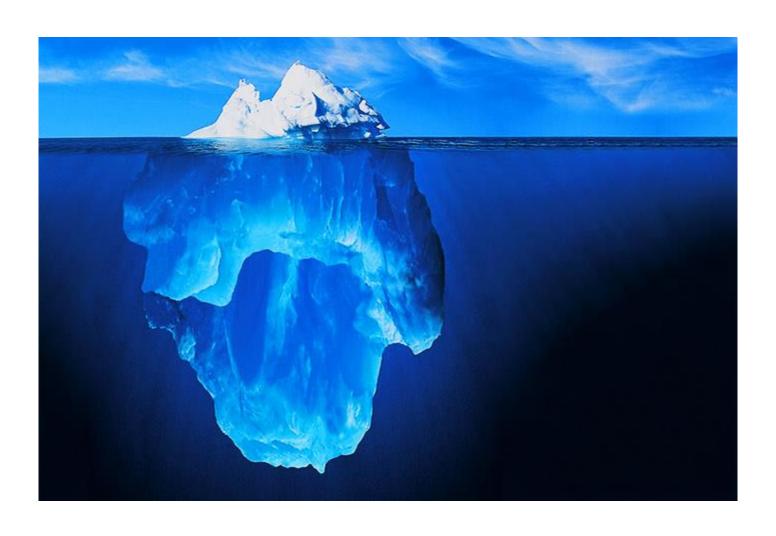


"Now! ... That should clear up a few things around here!"

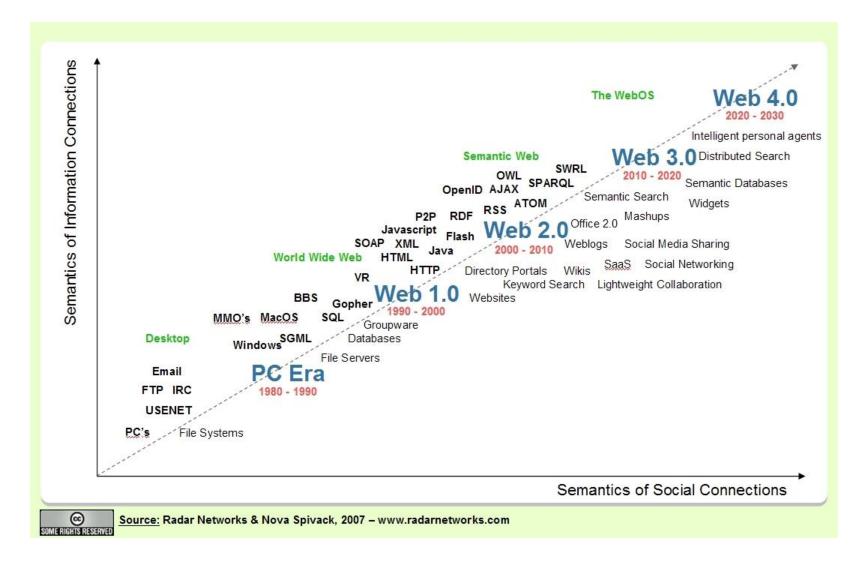
THE SEMANTIC WEB IS DATA INTEROPERABILITY?



THE SEMANTIC WEB IS SURFACING THE DEEP WEB?



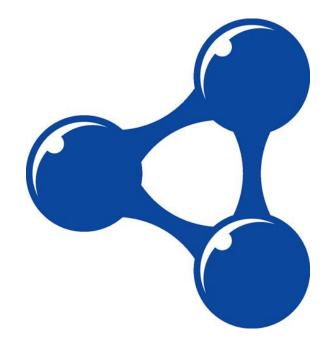
THE SEMANTIC WEB IS WEB 3.0?



The Semantic Web is ...

... A STACK OF WEB STANDARDS?









THE SEMANTIC WEB IS A RESEARCH AREA?



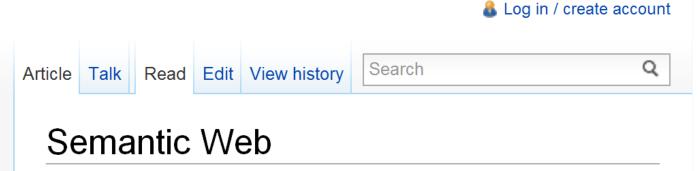
THE SEMANTIC WEB IS SOMETHING BAD?



Tom Coates; See talk: https://vimeo.com/21829271

THE SEMANTIC WEB IS A BUZZWORD?





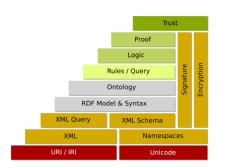


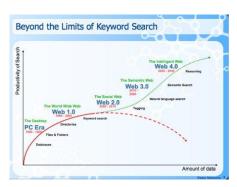
From Wikipedia, the free encyclopedia

Categories: Buzzwords | Semantic Web | Web services

THE SEMANTIC WEB IS...









Semantic Web

Categories: Buzzwords | Semantic Web | Web services

WIKIPEDIA

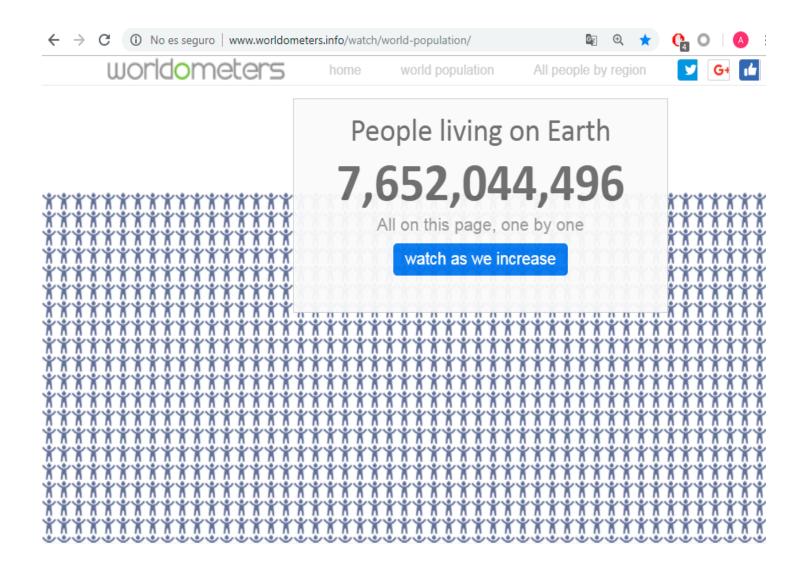


WHAT'S WRONG WITH THE CURRENT WEB?

THE CURRENT WEB IS FANTASTIC!

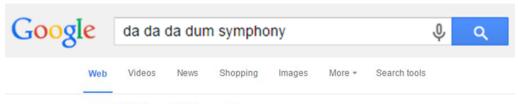


THE CURRENT WEB IS FANTASTIC!



Google is also pretty great

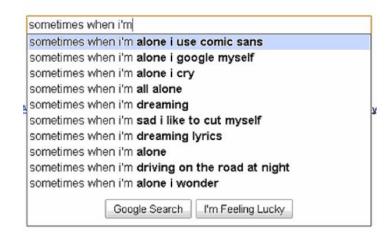




About 107,000 results (0.36 seconds)







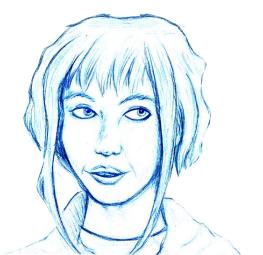
PROBLEM WITH THE CURRENT WEB: LITERATURE VETERANS

Doing a report for university ...



Wants to find all:

- Nobel Prize winners in Literature
- Who fought in at least one war
- The year they won the prize
- And the year the war(s) started



... how would you solve this?

LOTS OF WIKIPEDIA TABS ...





THE LAUREATE ALGORITHM (29 YEARS ON)



```
results := Ø
```

for all nobel-lit-winner in wiki-list

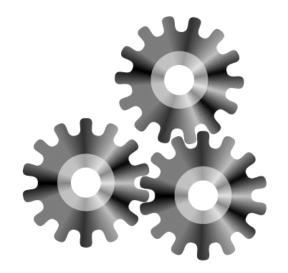
```
year-prize := year of nobel-lit-winner
wars := search "war", "conflict", "battle" in nobel-lit-winner
for all war in wars
    year-war := starting year of war
    add nobel-lit-winner, year-prize, war, year-war to results
end
```

end return results

THE LAUREATE ALGORITHM (58 YEARS ON?)



nobel-lit-winner-and-wars := magical-sem-web-results()
return nobel-lit-winner-and-wars

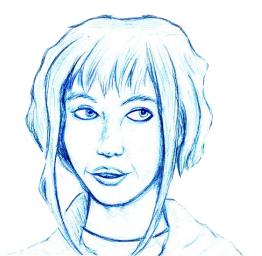


Doing a report for university ...



Wants to find all:

- Nobel Prize winners in Literature
- Who fought in at least one war
- The <u>year they won the prize</u>
- And the year the war(s) started



... why is this query hard on the current Web?

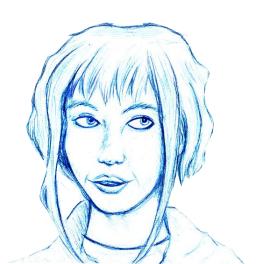
PROBLEM WITH THE CURRENT WEB: DESSERTS

GOING TO DINNER PARTY ...



Looking for a recipe:

- Must be a dessert
- Must be <u>citrus-free</u>
- Must be <u>gluten-free</u>
- Ingredients available in <u>local</u> supermarket(s)
- Cheap (student)
- Must be <u>delicious</u>



USING THE CURRENT WEB ...





THE DESSERT ALGORITHM (29 YEARS ON)



```
candidates := \emptyset
for all recipe-site in google-results
  for all dessert-recipe in recipe-site
     if dessert-recipe type looks-delicious
        suitable := true
        for all ingredient in dessert-recipe
           if searchNutrition(ingredient) type wheat or lemon or lime ...
               suitable := false
           else if searchShops(ingredient) type <u>null</u> or <u>expensive</u>
              suitable := false
           end
        end
         if suitable add dessert-recipe to candidates end
     end
  end
end
```

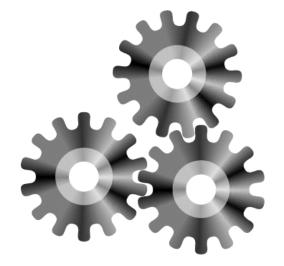
return candidates

THE DESSERT ALGORITHM (58 YEARS ON?)



```
candidates := Ø

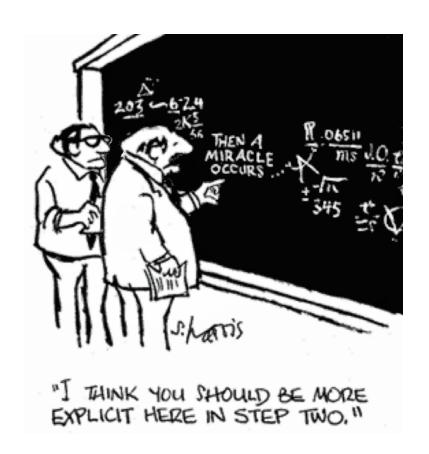
for all safe-cheap-local-dessert-recipe in magical-sem-web-results()
   if safe-cheap-local-dessert-recipe type looks-delicious
      add safe-cheap-local-dessert-recipe to candidates end
   end
end
return candidates
```

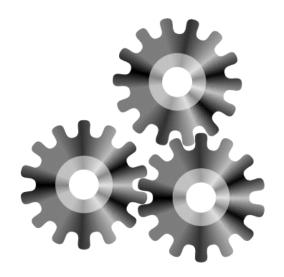


THE DESSERT ALGORITHM (58 YEARS ON?)



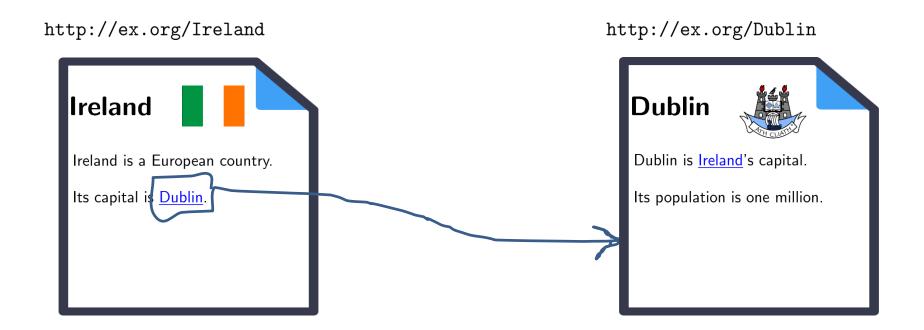
for all safe-cheap-local-dessert-recipe in <u>magical-sem-web-results()</u>





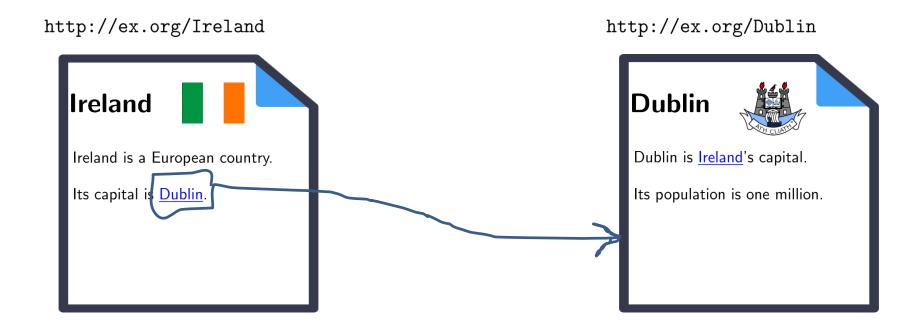
SO WHAT'S THE PROBLEM ...

THE CURRENT WEB IS DOCUMENT-CENTRIC

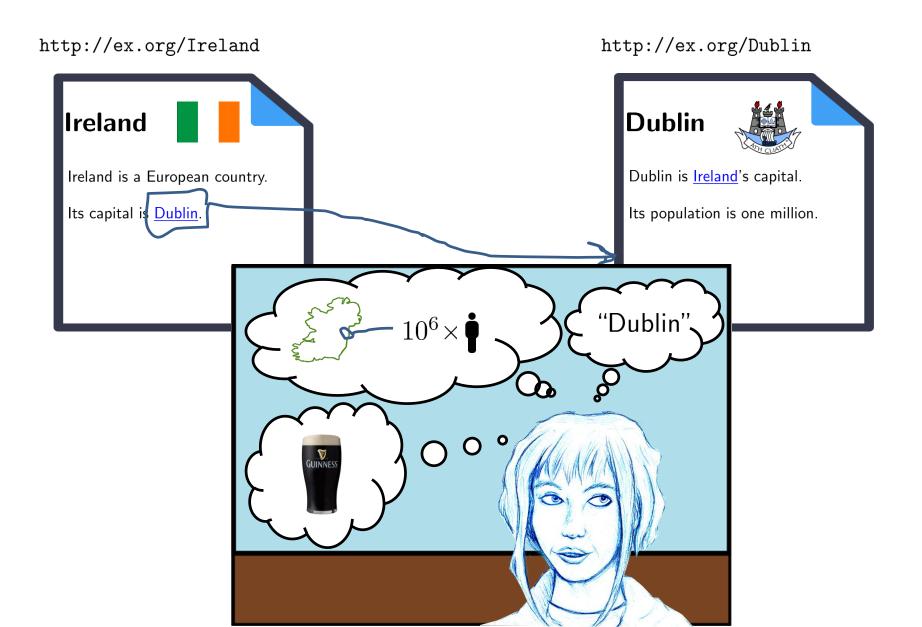


<html>

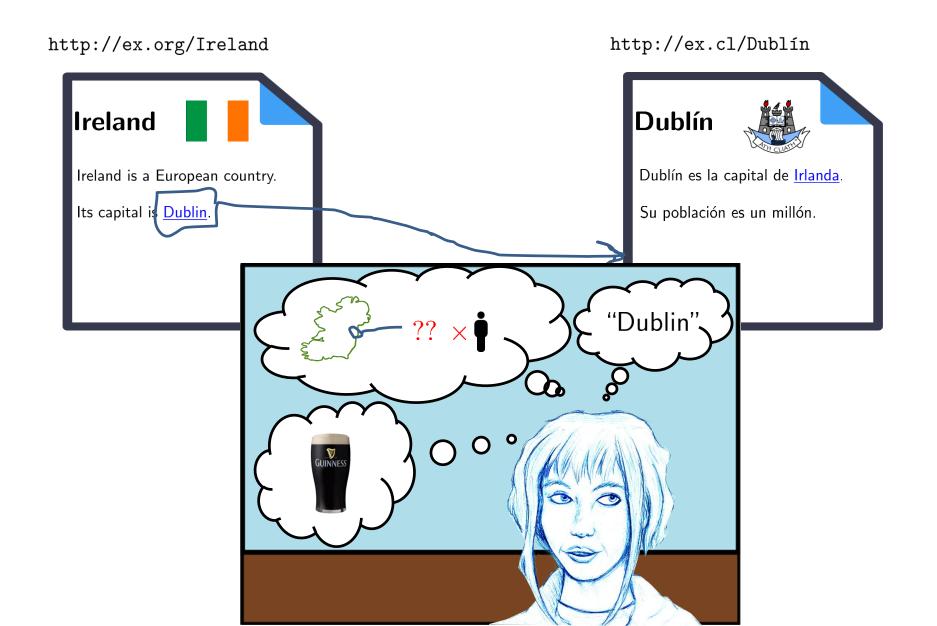
THE CURRENT WEB IS DOCUMENT-CENTRIC



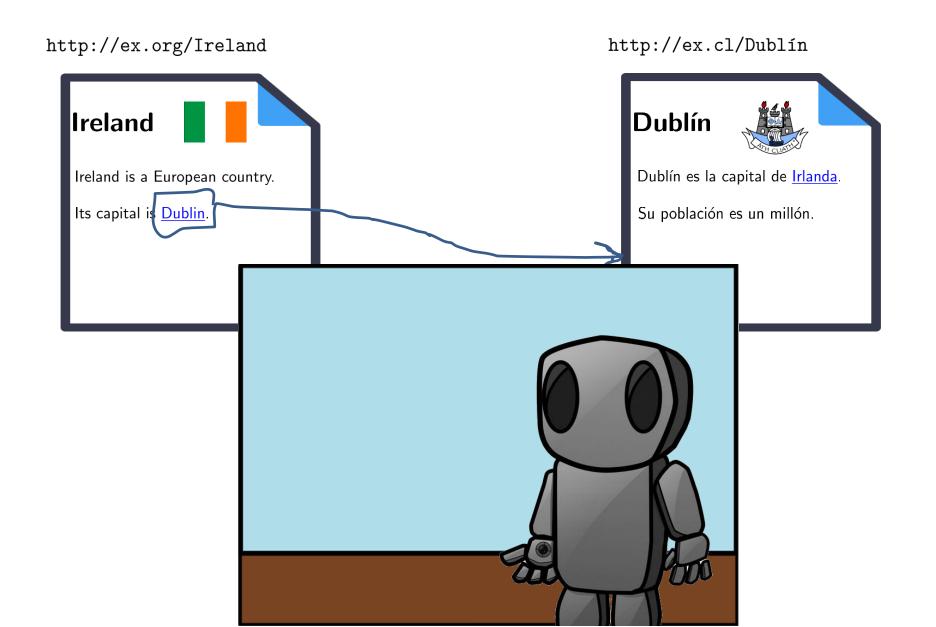
(Most of it) Makes sense to humans



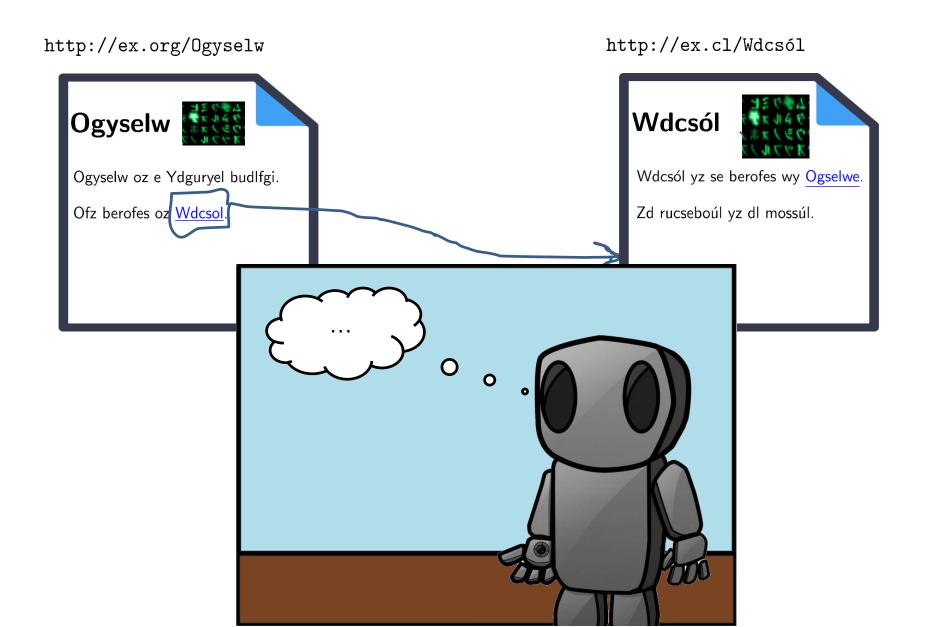
... ASSUMING THEY SPEAK THE LANGUAGE



EVEN WORSE FOR MACHINES

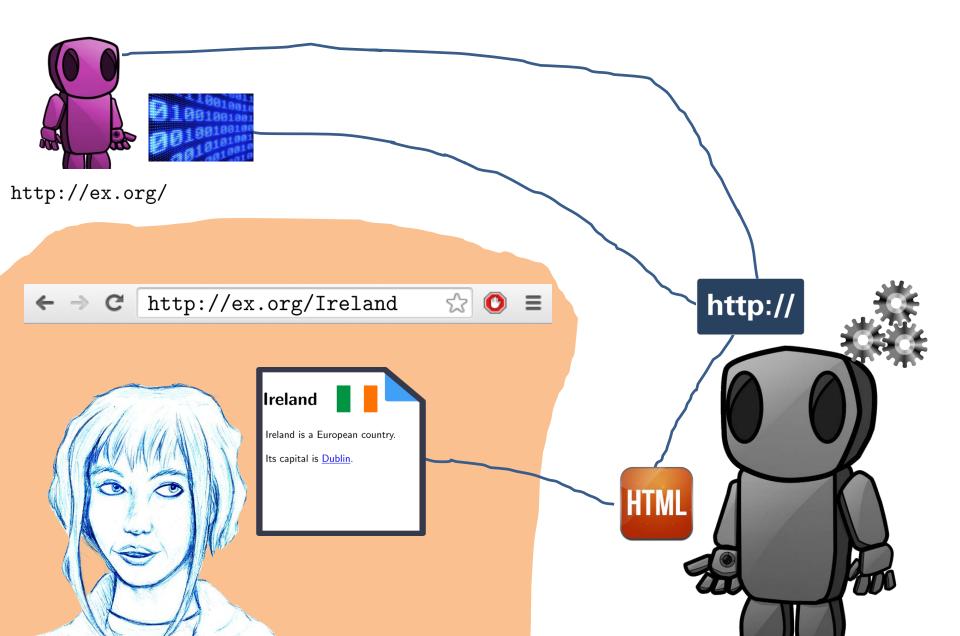


YTYL NUGZY PUG MEBHOLYZ

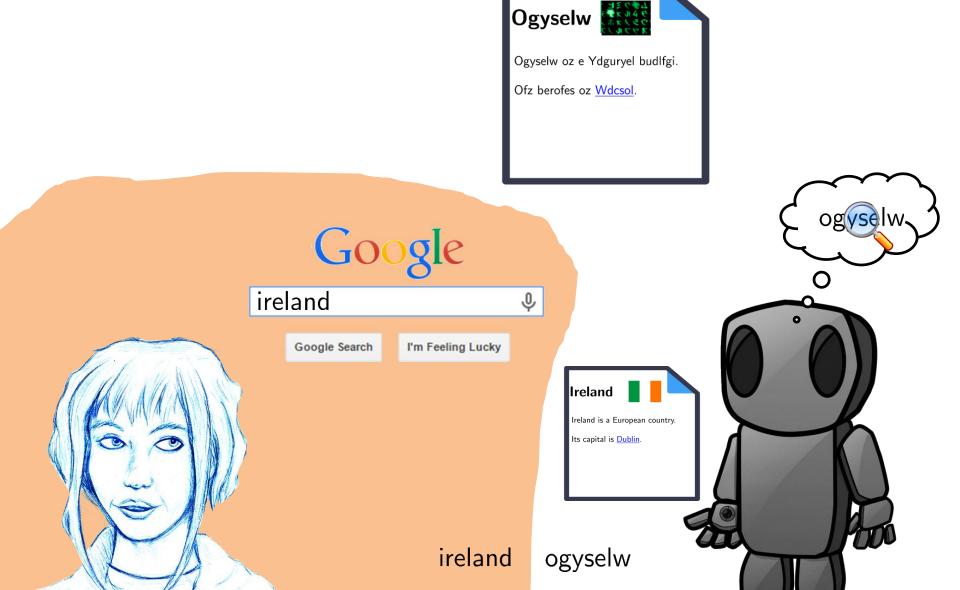


SO WHAT'S THE PROBLEM ...

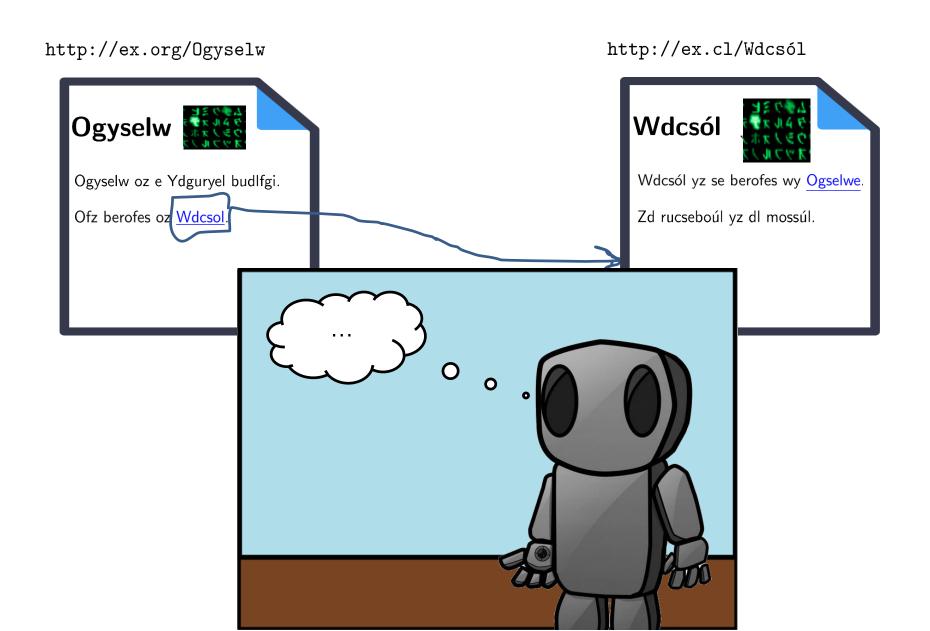
What machines can do: Fetch documents



What machines can do: Find documents

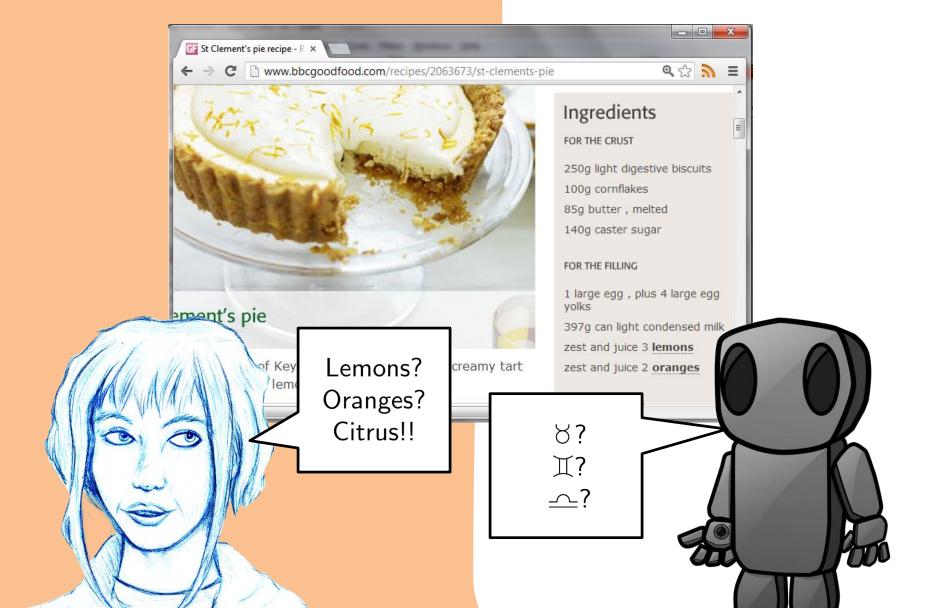


What machines cannot do: Combine sources

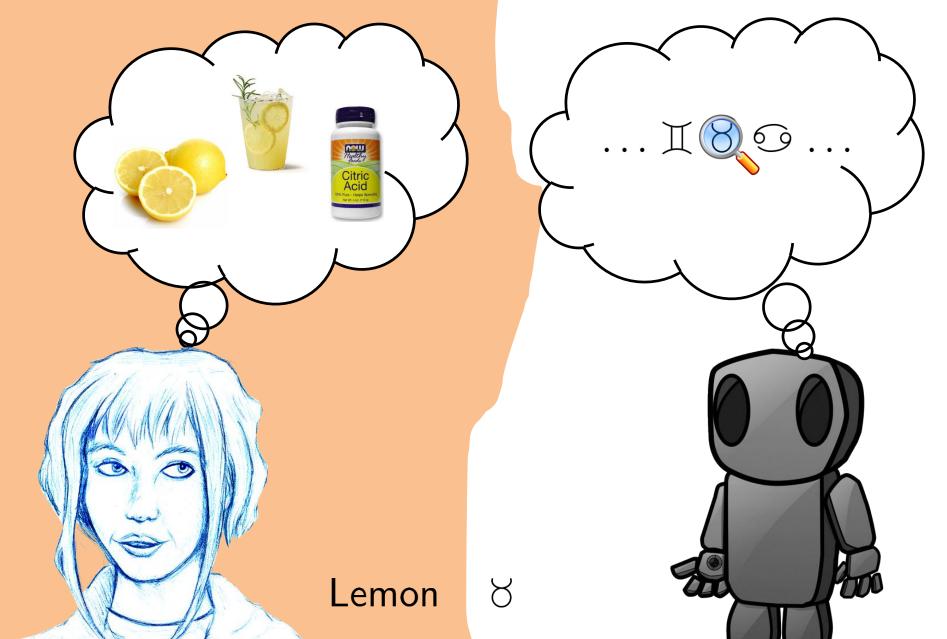


TOWARDS A SEMANTIC WEB

LET'S LOOK AT THIS RECIPE ...

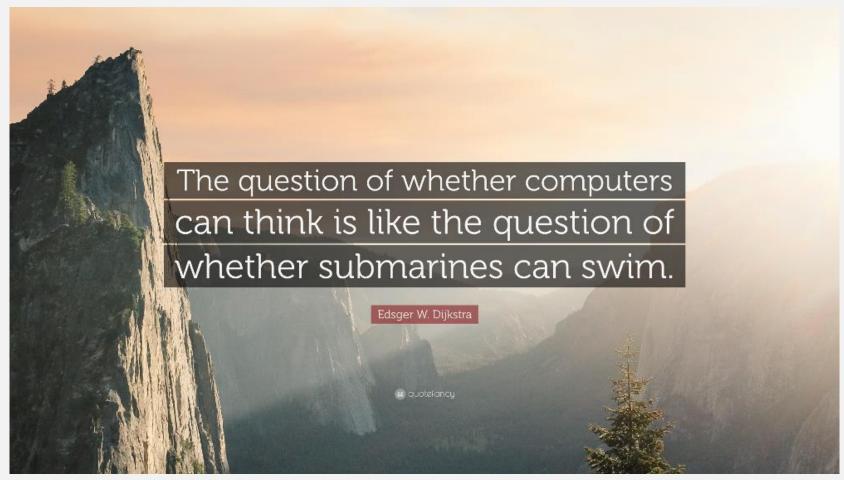


CAN MACHINES BE TAUGHT TO "UNDERSTAND"?



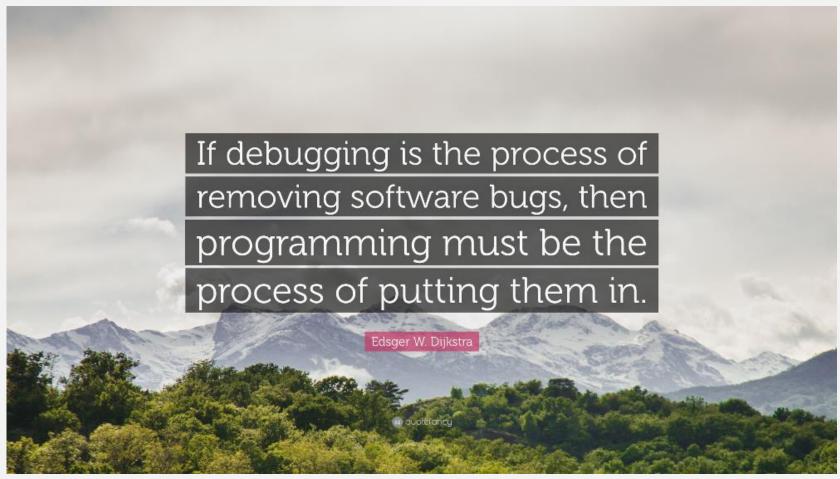
Edsger W. Dijkstra





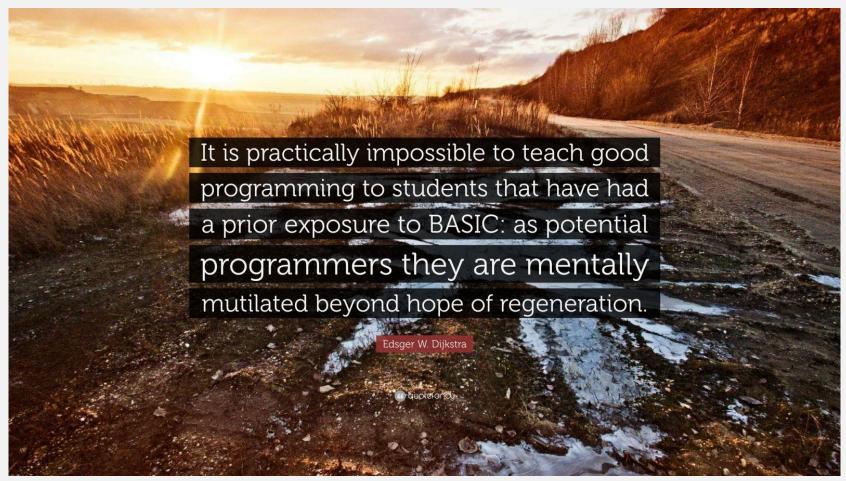
Edsger W. Dijkstra





EDSGER W. DIJKSTRA





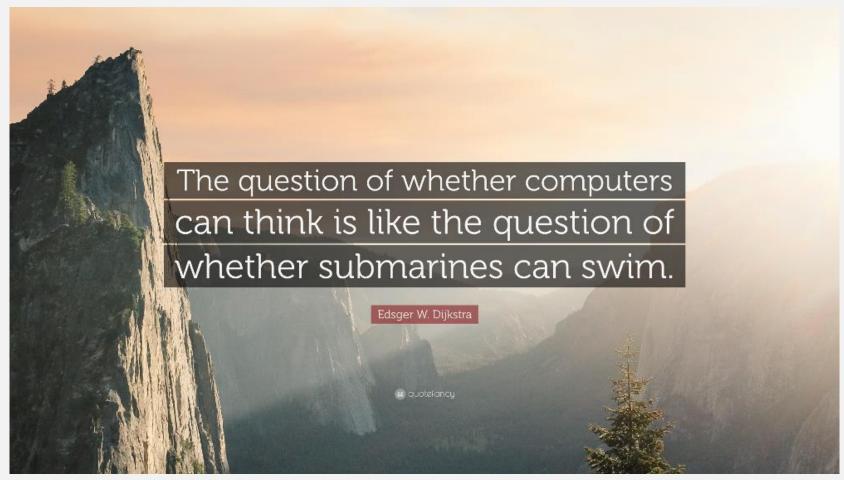
Edsger W. Dijkstra





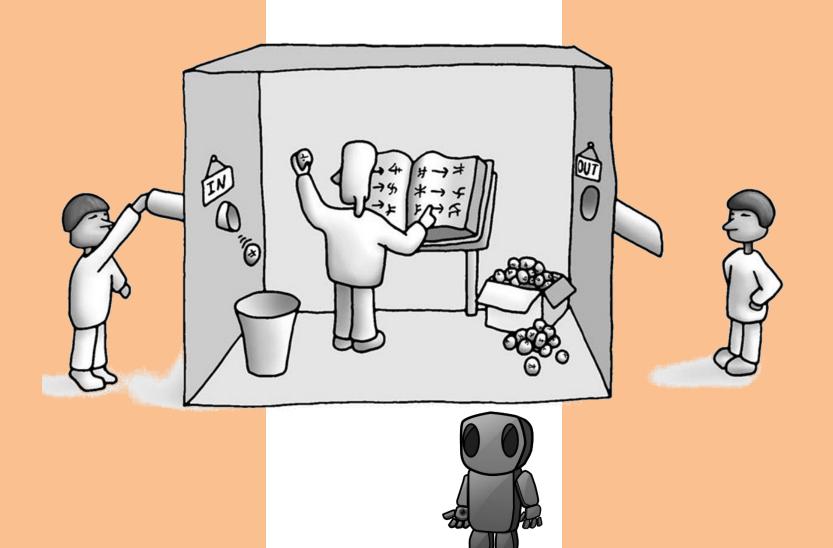
Edsger W. Dijkstra





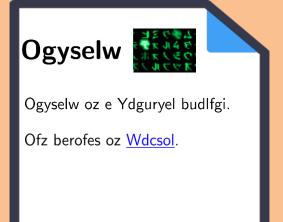
SEARLE'S CHINESE ROOM

SEARLE'S CHINESE ROOM ...



Searle's Chinese Room (Natural Language)

http://ex.org/Ogyselw



http://ex.org/Wdcsol

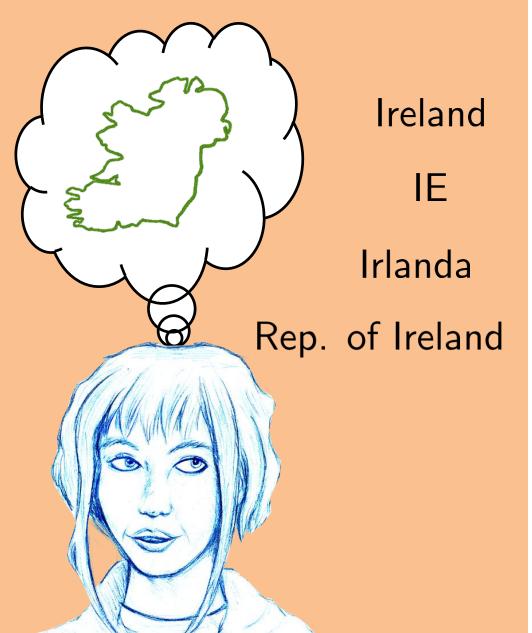


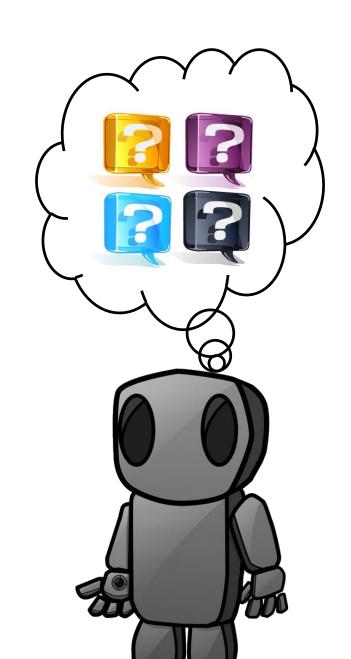
INPUT: "Nhef oz fhy rurdsefoul up fhy berofes up Ogyselw?"

... what should the output be?

OUTPUT: "uly mossoul"

MULTIPLE NAMES, ONE THING ...





ONE NAME, MULTIPLE THINGS ...





MULTIPLE WAYS TO SAY THE SAME THING ...

Dublin's population is one million.

Dublin has a population of one million.

Dublin's population is 1,000,000.

Dublin has 1,000,000 inhabitants.

One million people live in Dublin.

[Dublin] Its population is one million.

La población de Dublín es un millón.





MULTIPLE MEANINGS FOR THE SAME SAYING ...

Sherlock saw the man using binoculars.





NOT SAYING WHAT IS MEANT ...

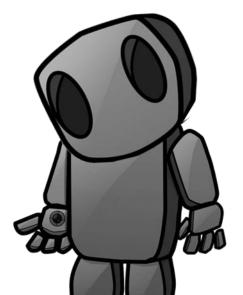
Fred está arriba la pelota.



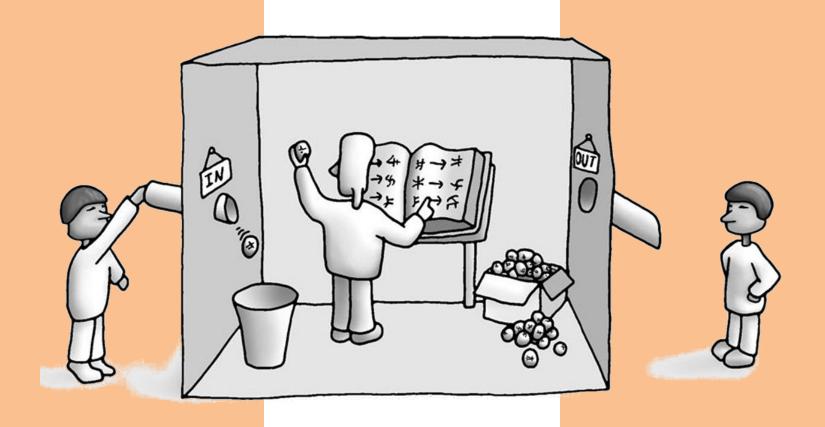


THE SEMANTIC GAP



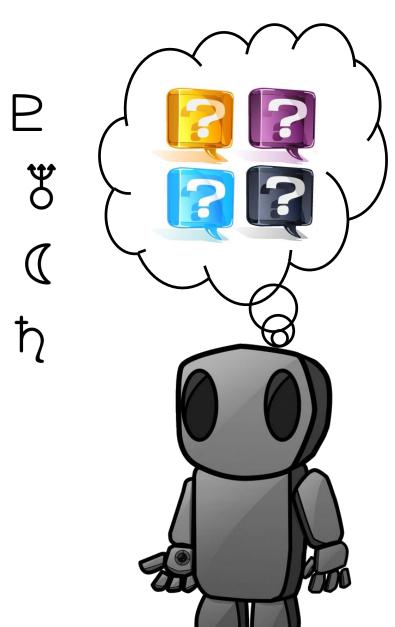


WHAT IF WE COULD "STRUCTURE" EVERYTHING ...



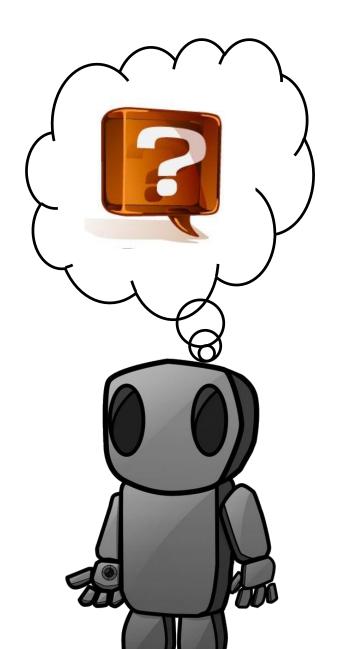
ONE SYMBOL, ONE MEANING ...





ONE SYMBOL, ONE MEANING ...





ONE (SIMPLE) WAY TO SAY ONE THING ...

Dublin's population is one million.

Dublin has a population of one million. Dublin's population is 1,000,000.

Dublin has 1,000,000 inhabitants.

One million people live in Dublin.

[Dublin] Its population is one million.

La población de Dublín es un millón.

(Dublin, population, 1000000)

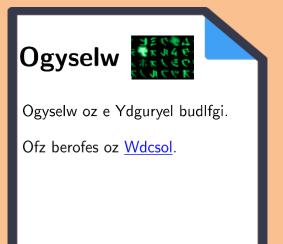
(♥,♂,1000000)





Searle's Chinese Room (Natural Language)

http://ex.org/Ogyselw



http://ex.org/Wdcsol



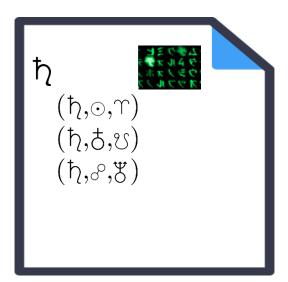
INPUT: "Nhef oz fhy rurdsefoul up fhy berofes up Ogyselw?"

... what should the output be?

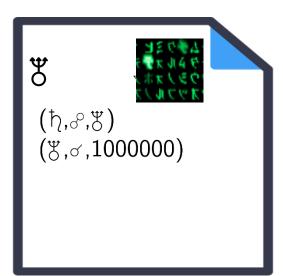
OUTPUT: "uly mossoul"

SEARLE'S CHINESE ROOM (SYMBOLIC)

http://ex.org/h



http://ex.org/∜



INPUT: " $(\uparrow, \varnothing, x), (x, \varnothing, y)$?"

... what should the output be?

Output: $\{(x \mapsto y, y \mapsto 1000000)\}$



SEARLE'S CHINESE ROOM (SYMBOLIC)

http://ex.org/Ireland



http://ex.org/Dublin



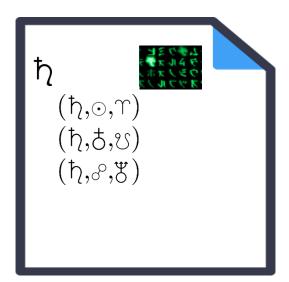
INPUT: "(Ireland, capital, x), (x, population, y)?"

... what should the output be?

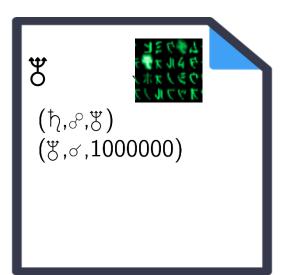
Output: $\{(x \mapsto \mathsf{Dublin}, y \mapsto 1000000)\}$



http://ex.org/h



http://ex.org/∜



INPUT: " (x, \odot, y) ?"

... what should the output be?

Output: $\{(x \mapsto h, y \mapsto \Upsilon)\}$



http://ex.org/Ireland



http://ex.org/Dublin



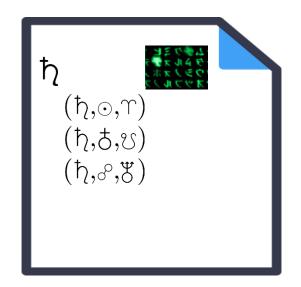
INPUT: "(x, partOf, y)?"

... what should the output be?

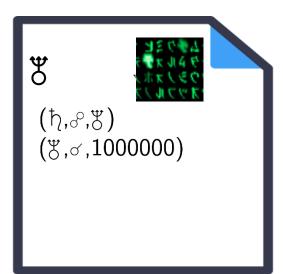
OUTPUT: $\{(x \mapsto \mathsf{Ireland}, y \mapsto \mathsf{Europe})\}\$ $(x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Ireland})$?



http://ex.org/\u03b7



http://ex.org/∜



Rule: " $(b, \beta, a) \rightarrow (a, \odot, b)$ "

Input: " (x, \odot, y) ?"

... what should the output be?

Output: $\{(x \mapsto h, y \mapsto \Upsilon), (x \mapsto f, y \mapsto h)\}$



http://ex.org/Ireland



http://ex.org/Dublin



RULE: " $(b, \mathsf{capital}, a) \to (a, \mathsf{partOf}, b)$ "

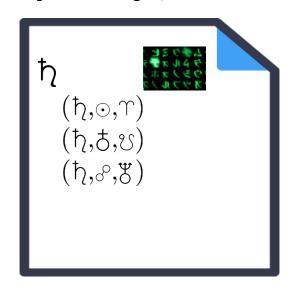
INPUT: "(x, partOf, y)?"

... what should the output be?

```
OUTPUT: \{(x \mapsto \mathsf{Ireland}, y \mapsto \mathsf{Europe}), (x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Ireland})\}
(x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Europe})?
```



http://ex.org/h



http://ex.org/∜

ψ (ħ,Φ,Ψ) (♥,σ,1000000)

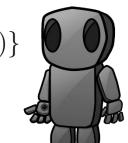
Rules:

$$"(b, \circ, a) o (a, \odot, b)"$$
 $"(a, \odot, b), (b, \odot, c) o (a, \odot, c)"$

INPUT: " (x, \odot, y) ?"

... what should the output be?

Output: $\{(x \mapsto h, y \mapsto \Upsilon), (x \mapsto f, y \mapsto h), (x \mapsto f, y \mapsto \Upsilon)\}$



http://ex.org/Ireland

Ireland



(Ireland, part Of, Europe) (Ireland, isA, Country) (Ireland, capital, Dublin)

http://ex.org/Dublin

Dublin



(Ireland, capital, Dublin) (Dublin, population, 1000000)

Rules:

```
"(b, \mathsf{capital}, a) \to (a, \mathsf{partOf}, b)"
"(a, \mathsf{partOf}, b), (b, \mathsf{partOf}, c) \to (a, \mathsf{partOf}, c)"
```

INPUT: "(x, partOf, y)?"

... what should the output be?

OUTPUT:
$$\{(x \mapsto \mathsf{Ireland}, y \mapsto \mathsf{Europe}), (x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Ireland}), (x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Europe})\}$$



SEMANTIC WEB: DATA → RULES → QUERY → OUTPUT*

DATA:





```
Rules: "(b, \mathsf{capital}, a) \to (a, \mathsf{partOf}, b)" "(a, \mathsf{partOf}, b), (b, \mathsf{partOf}, c) \to (a, \mathsf{partOf}, c)"
```

QUERY: "(x, partOf, y)?"

```
OUTPUT: \{(x \mapsto \mathsf{Ireland}, y \mapsto \mathsf{Europe}), \ (x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Ireland}), \ (x \mapsto \mathsf{Dublin}, y \mapsto \mathsf{Europe})\}
```



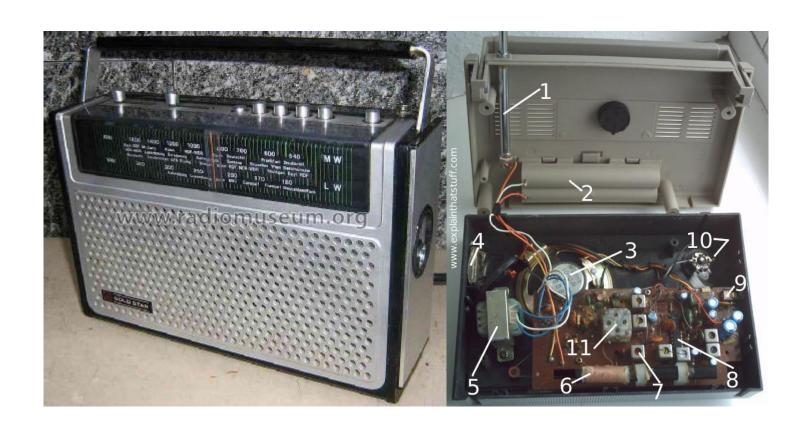
THE SEMANTIC WEB NOW?

THE SEMANTIC WEB IS NOW ABOUT 18 YEARS OLD



... so where is it then?

HIDDEN WITHIN THE WEB ...

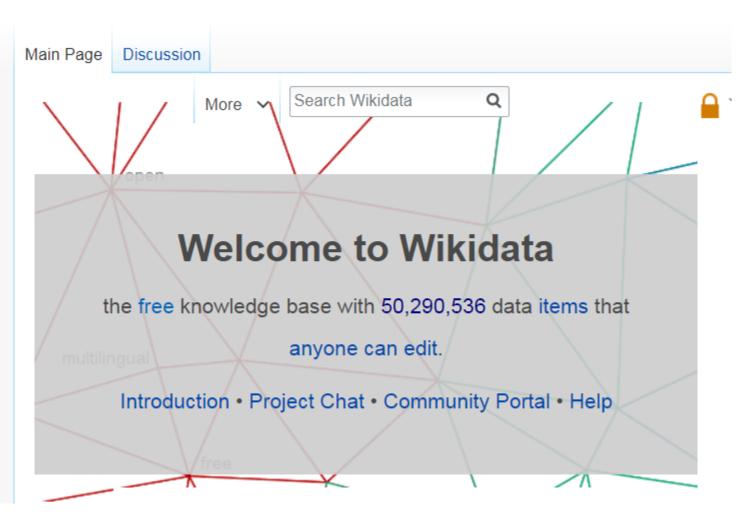


WIKIDATA: A WIKIPEDIA FOR DATA

WHAT IS WIKIDATA?



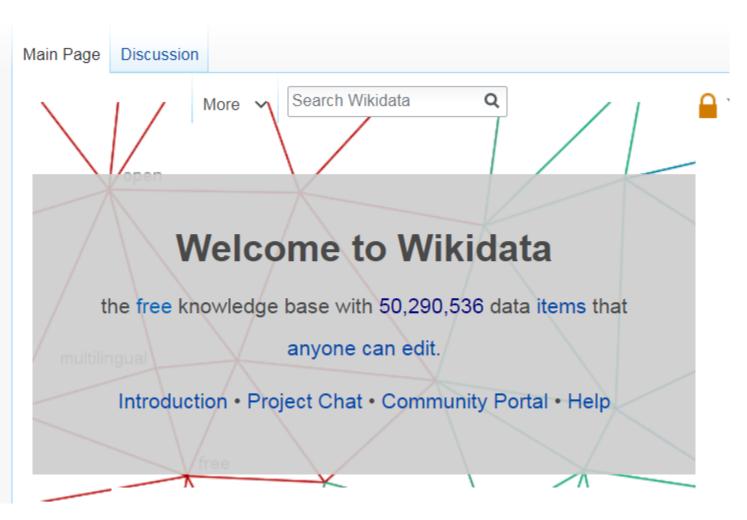
Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby
Help
Donate



WHY IS WIKIDATA?

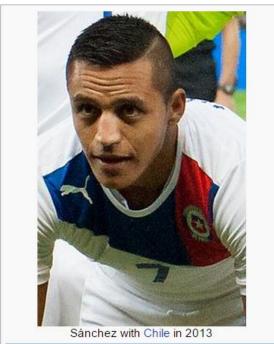


Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby
Help
Donate



PROBLEM 1: DIFFERENT LANGUAGE VERSIONS MANUALLY EDITED BY USERS

Alexis Sánchez



Personal information

Full name Alexis Alejandro Sánchez

Sánchez[1]

Date of birth 19 December 1988 (age 26)[1][2]

Place of birth Tocopilla, Chile [3][2]

| | National team [‡] | | |
|-------|----------------------------|----|------|
| 2007 | Chile U20 | 12 | (2) |
| 2006- | Chile | 82 | (26) |



| Carrera internacional | | | | | | |
|-----------------------|---------|--|--|--|--|--|
| Selección Lhile | | | | | | |
| Part. (goles) | 82 (26) | | | | | |
| Debut 2006 | | | | | | |



| | Nationalmannschaf | t ² |
|-------|-------------------|----------------|
| 2007 | Chile U-20 | |
| 2006– | Chile | 76 (25) |

PROBLEM 2: COMPLEX LISTS OF THINGS MANUALLY EDITED BY USERS





Most capped players [edit]

From Wikipedia, the free encyclopedia

As of September 1, 2016

Players in **bold** are still active, at least at club level.

Top goalscorers [edit]

As of September 1, 2016

Players in **bold** are still active, at least at club level.

| # | Name | International Career | Caps | Goa | ıls# | Name | International Career | Goals | Caps |
|----|-------------------|-------------------------|------|-----|------|-------------------|-------------------------|-------|------|
| 1. | Claudio Bravo | 2004 – | 106 | 0 | 1. | Marcelo Salas | 1994–2007 | 37 | 70 |
| 2. | Alexis Sánchez | 2006 – | 102 | 34 | ŀ | Iván Zamorano | 1987–2001 | 34 | 69 |
| 3. | Gary Medel | 2007 – | 96 | 7 | 2. | Alexis | | | |
| 4. | Gonzalo Jara | 2006 – | 95 | 3 | | Sánchez (list) | 2006 – | 34 | 102 |
| | | | | | | Educated a | | | |

ALEXIS SCORES A GOAL ...

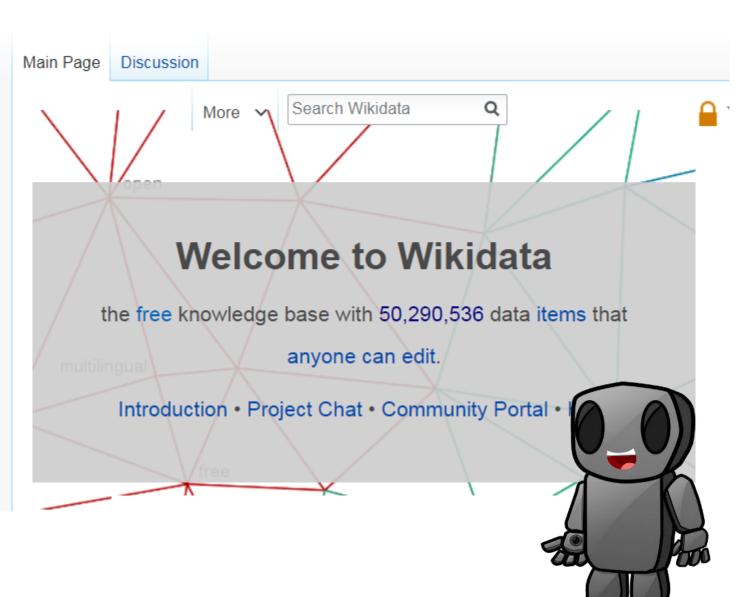


Now an army of human editors has to manually update a bunch of articles: different languages, lists, ...

SOLUTION: WIKIDATA



Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby
Help
Donate





Main page

Community portal

Project chat Create a new item

Item by title

Recent changes

Random item

Nearby

Help

Donate

Print/export

Create a book

Download as PDF

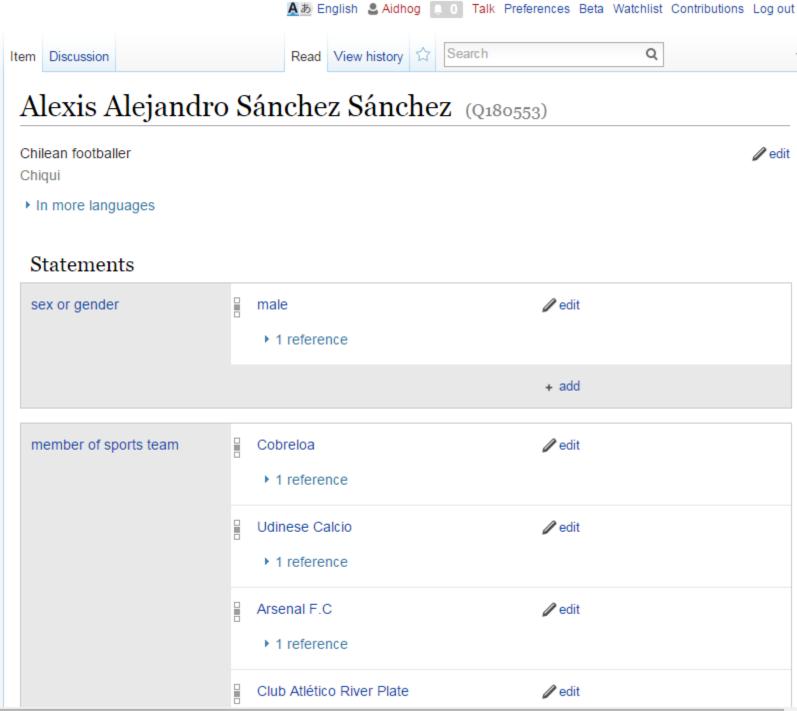
Printable version

Tools

What links here Related changes Special pages Permanent link

Page information Concept URI

Cite this page





Página principal

Portal de la comunidad

Café

Crear un elemento nuevo

Elemento por título

Cambios recientes

Elemento aleatorio

Cercanos Ayuda

Donaciones

Imprimir/exportar

Crear un libro

Descargar como PDF

Versión para imprimir

Herramientas

Lo que enlaza aquí Cambios relacionados Páginas especiales

Enlace permanente

Información de la página

Concept URI

Citar esta página



A b español 2 Aidhog 0

Leer Ver historial



Buscar

Discusión Preferencias Beta Lista de seguimiento Contribuciones Salir

Q

editar

Alexis Sánchez (Q180553)

futbolista chileno

Alexis Sánchez | AS9 | Alexis Alejandro Sánchez | AS7 | Alexis Alejandro Sánchez | Alexis Sánches | Alexis Sanchez | Alexis Sanches

En más idiomas

Declaraciones





Hauptseite

Gemeinschaftsportal

Forum

Ein neues Datenobjekt erstellen

Datenobjekte nach Titel

Letzte Änderungen

Zufälliges Datenobjekt

In der Nähe

Spenden

Hilfe

Drucken/exportieren

Buch erstellen

Als PDF herunterladen

Druckversion

Werkzeuge

Links auf diese Seite

Änderungen an verlinkten Seiten

Spezialseiten

Permanenter Link

Seiteninformationen

Konzept-URI

Seite zitieren



Aあ Deutsch 🚨 Aidhog 🔲 🕕

Lesen Versionsgeschichte 🗘



Suchen

Diskussion Einstellungen Beta Beobachtungsliste Beiträge Abmelden

Q

Alexis Sánchez (Q180553)

chilenischer Fußballspieler

Diskussion

bearbeiten

Alexis Alejandro Sánchez Sánchez | Alexis Sanchez | Alexis Alejandro Sanchez Sanchez

In weiteren Sprachen

Aussagen

Datenobjekt



| | | + mizaragen |
|-------------------------------------------|--------------------------------|-------------|
| | | |
| Mitglied von Sportmannschaft oder -verein | CD Cobreloa • 1 Fundstelle | |
| | Udinese Calcio ▶ 1 Fundstelle | |
| | FC Arsenal 1 Fundstelle | |
| | CA River Plate | |

Заглавная страница

Портал сообщества

Форум

Создание нового элемента

Элементы по заголовку

Свежие правки

Случайный элемент

Рядом Справка

Пожертвования

Печать/экспорт

Создать книгу

Скачать как PDF Версия для печати

Инструменты

Ссылки сюда

Связанные правки

Спецстраницы

Постоянная ссылка

Сведения о странице

URI концепта

Цитировать страницу

🛕 あ русский 🚨 Aidhog 🔲 🕕

Читать История 🛱

Поиск Q

Обсуждение Настройки Бета Список наблюдения Вклад Выйти

редактировать

редактировать

редактировать

+ добавить

редактировать

Санчес, Алексис (Q180553)

Описание не заполнено

Алексис Санчес

пол

Статья Обсуждение

На других языках

Утверждения

мужской

1 источник

член спортивной команды

Кобрелоа

1 источник

Удинезе

1 источник

Арсенал

1 источник

Ривер Плейт

редактировать

редактировать

WIKIDATA

സാമൂഹികകവാടം

പുതിയൊരു ഇനം സ്പ്ഷ്ലിക്കുക്

ഇനം തലക്കെട്ടനുസരിച്ച്

ഏതെങ്കിലും താൾ

പ്രധാന താൾ

Project chat

സമീപകാല

സമീപസ്ഥം

അച്ചടിയ്ക്കുക/ കയ്റ്റുമതി ചെയ്യുക

പുസ്തകം

PDF ആയി

പെകരണങ്ങൾ

സ്പ്ഷ്മിക്കുക

ഡൗൺലോഡ് ചെയ്യുക അച്ചടിരൂപം

സഹായം സംഭാവന

മാറ്റങ്ങൾ

അലക്സിസ് സാഞ്ചസ് (0180553)

വായിക്കുക നാൾവഴി കാണുക 😭

തിരയുക

വിവരണമൊന്നും നിർവചിച്ചിട്ടില്ല അപരനാമങ്ങളൊന്നും കണ്ടെത്താനായില്ല.

കൂടുതൽ ഭാഷകളിൽ

Statements

ലിംഗം 🖊 തിരുത്തുക പുരുഷൻ 1 സ്രോതസ്സ്

member of sports team *ഇംഗ്ലീഷ്*

Cobreloa *ഇംഗ്ലീഷ്*

🖊 തിരുത്തുക

🖊 തിരുത്തുക

+ ചേർക്കുക

Q

🥒 തിരുത്തുക

1 സ്രോതസ്സ്

1 സ്രോതസ്സ്

Udinese Calcio *ಅಂಭಿ\ഷ്* 1 സ്രോതസ്സ്

ആഴ്സണൽ എഫ്.സി.

Club Atlético River Plate ഇംഗ്ലീഷ്

🖊 തിരുത്തുക

🖊 തിരാത്താക

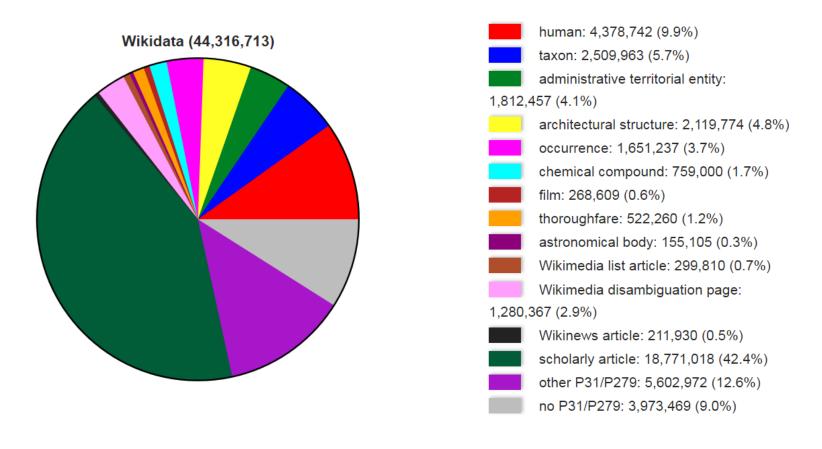
പ്രത്യേക താളുകൾ സ്ഥിരംകണ്ണി

അനുബന്ധ മാറ്റ്ങ്ങൾ

താദിന്റെ

ഈ താളിലേക്കുള്ള കണ്ണികൾ

What does Wikidata describe?



• 19,281 active users

Use-case: Info-Boxes



Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help About Wikipedia Community portal Recent changes Contact page Article Talk

Atacama Pathfinder Experiment

From Wikipedia, the free encyclopedia

The Atacama Pathfinder Experiment (APEX) is a radio telescope 5,100 meters above sea level, at the Llano de Chajnantor Observatory in the Atacama desert in northern Chile, 50 km east of San Pedro de Atacama built and operated by 3 European research institutes. The main dish has a diameter of 12 m and consists of 264 aluminium panels with an average surface accuracy of 17 micrometres (rms). The telescope was officially inaugurated on September 25, 2005.

The APEX telescope is a modified ALMA (Atacama Large Millimeter Array) prototype antenna and is at the site of the ALMA observatory. APEX is designed to work at sub-millimetre wavelengths, in the 0.2 to 1.5 mm range — between infrared light and radio waves — and to find targets that ALMA will be able to study in greater detail. Submillimetre astronomy provides a window into the cold, dusty and distant Universe, but the faint signals from space are heavily absorbed by water vapour in the Earth's atmosphere. Chajnantor was chosen as the location for such a telescope because the region is one of the driest on the planet and is more than 750

Atacama Pathfinder Experiment



The APEX telescope

Observatory Llano de Chajnantor

Observatory P

Location(s) Atacama Desert, Chile

n

Organization Edit this at Wikidata

Observatory

Max Planck Institute for

Radio Astronomy

Onsala Space Observatory -

Altitude 5,100 m (16,700 ft)

Wavelength 0.2, 1.5 mm (1.50, 0.20 THz)

First light 2004

Telescope style Cassegrain reflector

Cosmic microwave background experiment Radio telescope

Use-case: Info-Boxes



Atacama Pathfinder Experiment



The APEX telescope

Llano de Chajnantor Observatory

Observatory P

Location(s) Atacama Desert, Chile

Coordinates 23°00′21″S 67°45′33″W

Organization Edit this at Wikidata

Observatory

Max Planck Institute for

Radio Astronomy

Onsala Space Observatory -

Altitude 5,100 m (16,700 ft)

Wavelength 0.2, 1.5 mm (1.50, 0.20 THz)

First light 2004

edit

edit

+ add reference

+ add reference

Telescope style Cassegrain reflector

Cosmic microwave background experiment

Radio telescope /



Special pages

Use-case: Quality Checks

List of all person who do not have an age between 0 and 130. Update: 22:21, 10 September 2018 (UTC)

| ltem ♦ | Birth ♦ | Death ♦ | Age + |
|----------------------------------------------------------|-------------|-------------|-------|
| Tuti Yusupova (Q1038827) | 1880-07-01 | 2015-03-28 | 134 |
| Karni Mata (Q1106783) | 1387-10-02 | 1538-03-23 | 150 |
| no label (Q11556831) | 1185-01-01 | 1392-01-01 | 207 |
| Minamoto no Chikayuki (Q11564306) | 1185-01-01 | 1333-01-01 | 148 |
| Bir Narayan Chaudhuri (Q11854281) | 1857-00-00 | 1998-04-20 | 141 |
| no label (Q12123094) | 1091-00-00 | 1228-00-00 | 137 |
| no label (Q12218744) | 1819-00-00 | 1954-00-00 | 135 |
| Maftei Pop (Q12734691) | 1804-01-01 | 1952-03-15 | 148 |
| Habib Miyan (Q1365575) | 1868-05-20 | 2008-08-19 | 140 |
| Egyō (Q1392070) | 805-01-01 | 1185-01-01 | 380 |
| Xu Xun (Q1428729) | 239-01-01 | 374-01-01 | 135 |
| Zaro Aga (Q148028) | 1777-01-01 | 1934-06-29 | 157 |
| Charles Étienne Guillaume Blandin de Chalain (Q15967550) | 1740-06-07 | 1958-01-01 | 217 |
| Javier Pereira (Q15999178) | 1789-00-00 | 1958-00-00 | 169 |
| Chen Jun (Q16077971) | 881-00-00 | 1324-00-00 | 443 |
| Jules Granier (Q16842647) | 1770-01-01 | 1906-04-07 | 136 |
| Genson (Q18115051) | 1700-01-01 | 1950-01-01 | 250 |
| Liutwin (Q18222784) | 1200-01-01 | 1350-01-01 | 150 |
| Salah (Q1827950) | -2068-00-00 | -1635-00-00 | 433 |

Use-case: Quality Checks

List of all person who do not have an age between 0 and 130. Update: 22:21, 10 September 2018 (UTC)

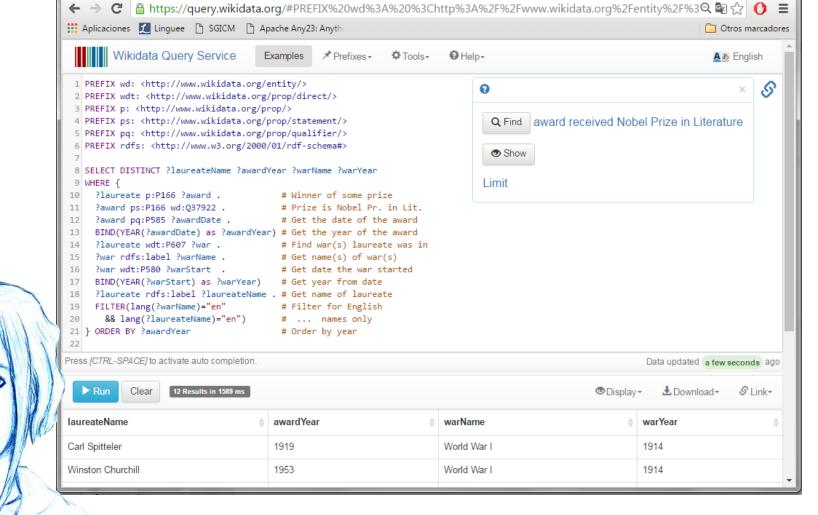
| Item | | Death ♦ | Age + |
|------------------------------------------------|-------------|-------------|-------|
| Tuti Yusupova (Q1038827) | 1880-07-01 | 2015-03-28 | 134 |
| Karni Mata (Q1106783) | 1387-10-02 | 1538-03-23 | 150 |
| no label (Q11556831) | 1185-01-01 | 1392-01-01 | 207 |
| Minamoto no Chikayuki (Q11564306) | 1185-01-01 | 1333-01-01 | 148 |
| Julius Fessler (Q55677113) | 1982-05-04 | 1937-10-27 | -45 |
| Max Hallbauer (Q55678777) | 1851-12-04 | 1818-10-08 | -34 |
| Ernst Keiter (Q55679293) | 1943-10-28 | 1907-10-30 | -36 |
| Franz Friedrich Theodor Steinhauer (Q55680095) | 1849-12-25 | 1822-12-21 | -28 |
| Johann Joseph Schoder (Q55680781) | 1918-12-07 | 1884-12-12 | -34 |
| Paul Kaspar Helbling (Q55681272) | 2012-06-05 | 2011-06-20 | -1 |
| Gustaf Bolinger (Q55683866) | 1988-12-19 | 1957-07-16 | -32 |
| Joseph Marcus Jaffé (Q55683961) | 1867-03-11 | 1841-04-12 | -26 |
| Johann Baptist Hau (Q55683983) | 1782-08-14 | 1758-12-26 | -24 |
| Wilhelm Klaubert (Q55684184) | 1557-08-17 | 1526-10-03 | -31 |
| Dieterich Johann Krüger (Q55684826) | 1742-06-23 | 1726-08-22 | -16 |
| Alfred Reichenbecher (Q55684894) | 1884-01-26 | 1664-03-05 | -220 |
| Johann Ludwig Winckler (Q55901640) | 1963-05-08 | 1767-08-08 | -196 |
| Johann Heremberck (Q55902890) | 1811-01-01 | 1489-01-01 | -322 |
| Salah (Q1827950) | -2068-00-00 | -1635-00-00 | 433 |

Use-case: Query Service

IIII Wikidata Query Service



Aîdan 🗆 🗆 X



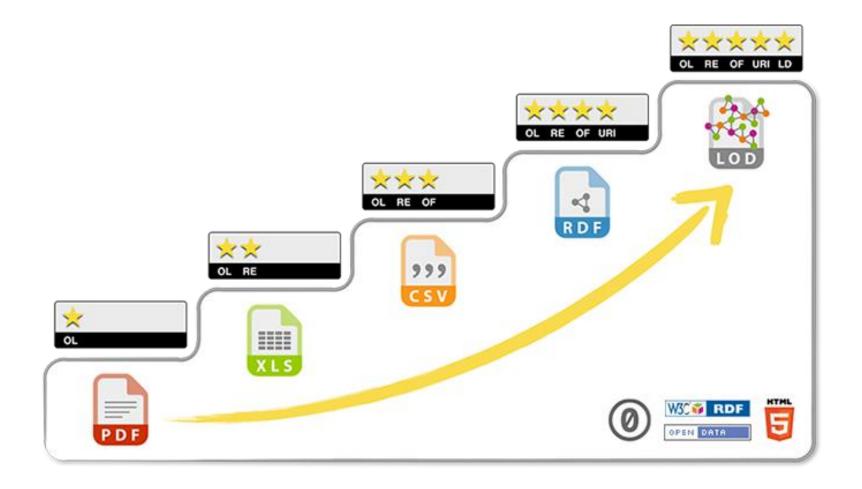
Use-case: Query Service



| IIII Wikidata Query Service × | | | | | | Aidan |
|-------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|------------------------------------------------|------|----------------|-------------|-------------------------------|
| | | | | | | ntity%2F%3 Q ᠍ ☆ 🚺 |
| Aplicaciones Linguee | SGICM 🖺 Apa | ache Any23: Anythi | | | | Otros marcado |
| 17 BIND(YEAR(?warStart) a ?laureate rdfs:label : 19 FILTER(lang(?warName)= 20 && lang(?laureateName) } ORDER BY ?awardYear 22 | as ?warYear) ?laureateName ="en" | # Get name of laureate # Filter for English | eu | | | |
| Press [CTRL-SPACE] to activate aut | to completion. | | | | | Data updated a few seconds ag |
| ► Run Clear 12 Result | s in 1589 ms | | | | ● Display * | ≛ Download → S Link → |
| laureateName | \$ | awardYear | | Name | ☆ | warYear |
| Carl Spitteler | | 1919 | Wor | ld War I | | 1914 |
| Winston Churchill | | 1953 | Wor | ld War I | | 1914 |
| Ernest Hemingway | | 1954 | Wor | ld War I | | 1914 |
| Ernest Hemingway | | 1954 | Wor | ld War II | , | 1939 |
| Jean-Paul Sartre | | 1964 | Alge | rian War | , | 1954 |
| Jean-Paul Sartre | | 1964 | Wor | ld War II | | 1939 |
| Heinrich Böll | | 1972 | Wor | ld War II | | 1939 |
| Eugenio Montale | | 1975 | Wor | ld War I | | 1914 |
| William Golding | | 1983 | Wor | ld War II | | 1939 |
| Claude Simon | | 1985 | Spa | nish Civil War | | 1936 |
| Camilo José Cela | | 1989 | Spa | nish Civil War | | 1936 |
| Günter Grass | | 1999 | Wor | ld War II | , | 1939 |

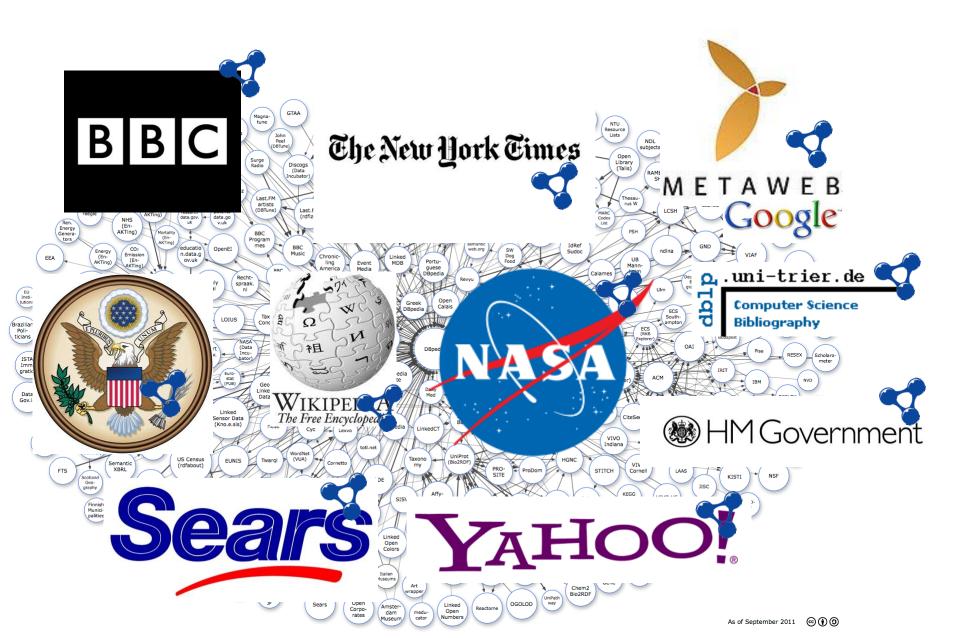
STANDARDS FOR OPEN DATA: LINKED OPEN DATA

HOW TO PUBLISH OPEN DATA?



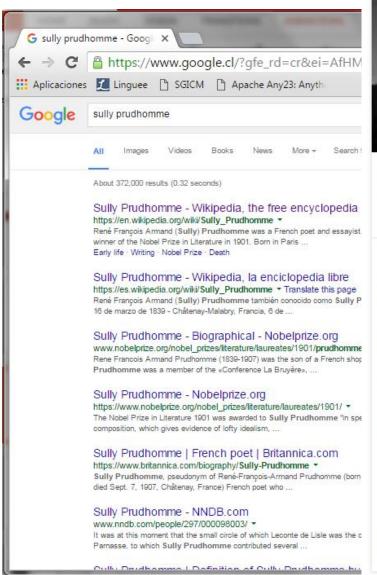
5-Star Linking Open Data Scheme

THE LINKED DATA CLOUD



GOOGLE'S KNOWLEDGE GRAPH

GOOGLE'S INFO-BOX





Sully Prudhomme



Poet

René François Armand Prudhomme was a French poet and essayist. He was the first ever winner of the Nobel Prize in Literature in 1901. Wikipedia

Born: March 16, 1839, Paris, France

Died: September 6, 1907, Châtenay-Malabry, France

Books: Les vaines tendresses

Awards: Nobel Prize in Literature

People also search for



Leconte de Lisle



Theodor Mommsen



Frédéric Mistral

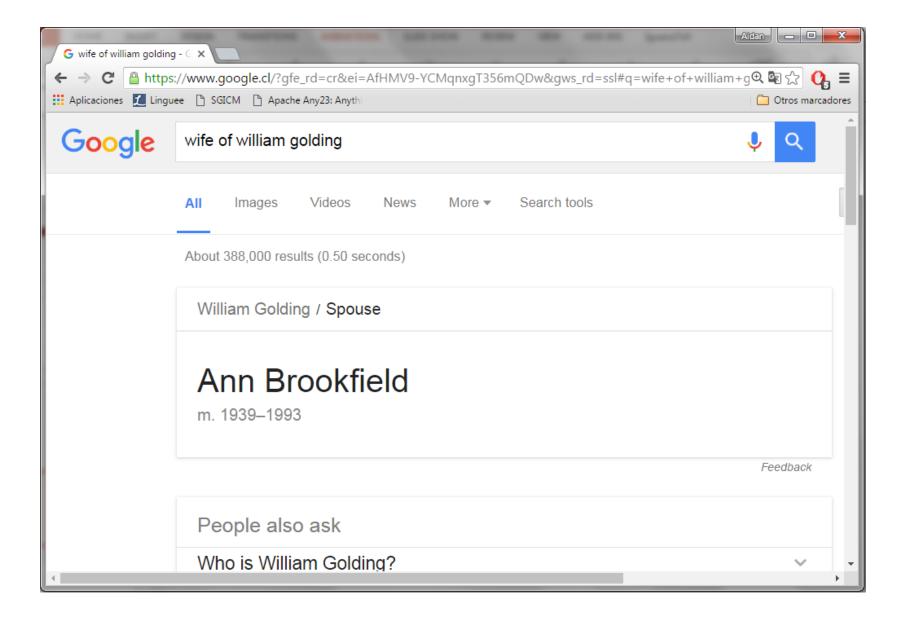


Paul Verlaine





GOOGLE'S SEMANTIC SEARCH



Using Semantic Web knowledge-bases

From Freebase to Wikidata: The Great Migration

Thomas Pellissier Tanon Google, San Francisco, USA thomas@pellissier-tanon.fr vrandecic@google.com

Denny Vrandečić Google, San Francisco, USA

Sebastian Schaffert Google, Zürich, Switzerland schaffert@google.com

Thomas Steiner Google, Hamburg, Germany tomac@google.com

Lydia Pintscher Wikimedia, Berlin, Germany lydia@pintscher.de

ABSTRACT

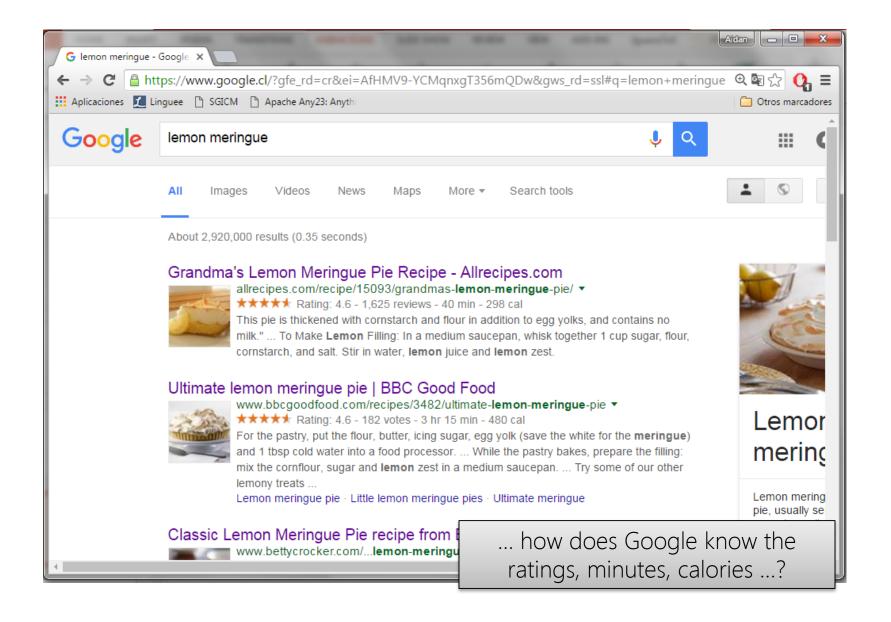
Collaborative knowledge bases that make their data freely available in a machine-readable form are central for the data strategy of many projects and organizations. The two major collaborative knowledge bases are Wikimedia's Wikidata and Google's Freebase. Due to the success of Wikidata. Google decided in 2014 to offer the content of Freebase to the Wikidata community. In this paper, we report on the ongoing transfer efforts and data mapping challenges, and provide an analysis of the effort so far. We describe the Primary Sources Tool, which aims to facilitate this and future data migrations. Throughout the migration, we have gained deep insights into both Wikidata and Freebase, and share and discuss detailed statistics on both knowledge bases.

One such collaborative knowledge base is Freebase, publicly launched by Metaweb in 2007 and acquired by Google in 2010. Another example is Wikidata, a collaborative knowledge base developed by Wikimedia Deutschland since 2012 and operated by the Wikimedia Foundation. Due to the success of Wikidata, Google announced in 2014 their intent to shut down Freebase and help the community with the transfer of Freebase content to Wikidata [10].

Moving data between two knowledge bases that do not share a similar design is usually a problematic task and requires the careful mapping between their structures. The migration from Freebase to Wikidata was no exception to this rule: we encountered a number of to-be-expected structural challenges. However, even more demanding was the cultural difference between the two involved communities.

GOOGLE'S RICH SNIPPETS

FANCY-LOOKING SEARCH RESULTS ...



Publishers add structured data!

```
  <span class="nutrition__label">kcal</span>
        <span class="nutrition__value" itemprop="calories">480</span>
        \ldots
```

- ✓ Publishers get more clicks on their results
- ✓ Google gets data to make fancy results

FACEBOOK'S OPEN GRAPH PROTOCOL

On Facebook ...

Facebook wants to know what pages linked in comments are about



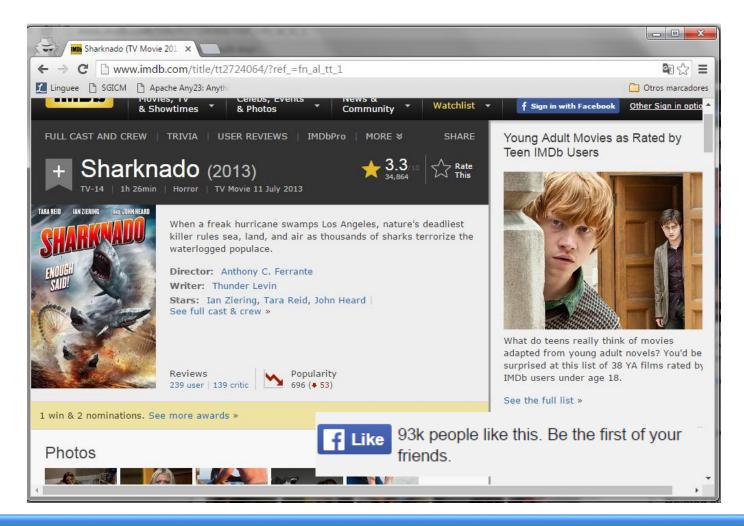
The latest trailer for "Game of Thrones" is out, and contains so many fantastic details that we don't even know where to begin. Take a look: http://imdb.to/1KXLPFJ



"Game of Thrones": The Wheel
"I'm not going to stop the wheel. I'm going to break the wheel."

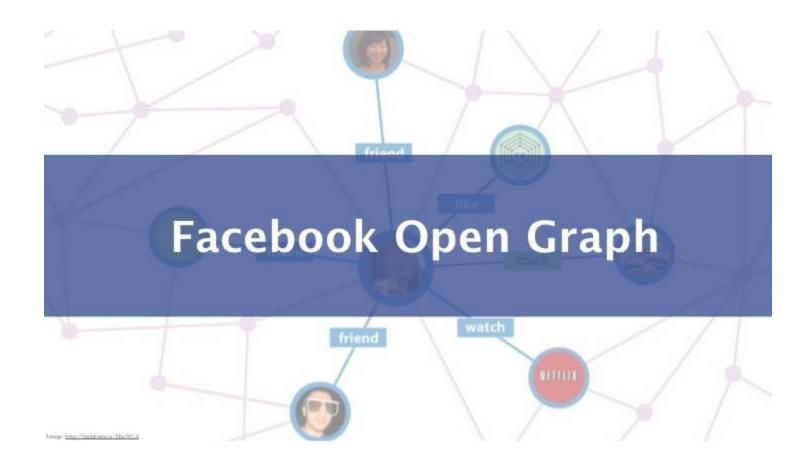
IMDB.COM

Elsewhere on the Web ...



Facebook wants to know what you like on other sites ... is it a movie, a recipe ...

FACEBOOK OPEN GRAPH PROTOCOL



ABOUT THE COURSE ...

TOPICS COVERED

- RDF (triple-based data model)
- RDFS/OWL (ontological languages)
- SPARQL (query language)
- Linked Data / Web of Data
- RDB2RDF (importing databases to Sem. Web.)

STRUCTURE OF THE COURSE

- Each week:
 - Class on Monday (learn concepts)
 - Lab on Wednesday (see concepts in practice)

Marking structure:

```
50% labs
```

two controls

15% project

5% reading group

OUTCOMES: LEARN ABOUT THE SEMANTIC WEB!

An ongoing research topic here in the DCC

Apply database, logic, AI, etc., to the Web

Mix of theory and practical exercises

The future of the Web?

